Systems Biology of Metabolism: A Driver for Developing Personalized and Precision Medicine

Nielsen, Jens

Published in:
Cell Metabolism

Link to article, DOI:
10.1016/j.cmet.2017.02.002

Publication date:
2017

Document Version
Publisher's PDF, also known as Version of record

CellPress

Cell Metabolism
Perspective

Systems Biology of Metabolism: A Driver for Developing Personalized and Precision Medicine

Jens Nielsen[1,2,3,*]
[1]Department of Biology and Biological Engineering, Chalmers University of Technology, SE41128 Gothenburg, Sweden
[2]Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, DK2800 Lyngby, Denmark
[3]Science for Life Laboratory, Royal Institute of Technology, SE17121 Stockholm, Sweden
*Correspondence: nielsenj@chalmers.se
http://dx.doi.org/10.1016/j.cmet.2017.02.002

Systems biology uses mathematical models to analyze large datasets and simulate system behavior. It enables integrative analysis of different types of data and can thereby provide new insight into complex biological systems. Here will be discussed the challenges of using systems medicine for advancing the development of personalized and precision medicine to treat metabolic diseases like insulin resistance, obesity, NAFLD, NASH, and cancer. It will be illustrated how the concept of genome-scale metabolic models can be used for integrative analysis of big data with the objective of identifying novel biomarkers that are foundational for personalized and precision medicine.

Introduction

Healthcare costs are rapidly increasing in the developing countries, and in 2011 the total healthcare spending in the United States accounted for about 18% of its GDP (WHO, 2011), a 63% inflation-adjusted increase since 1997 (Pfuntner et al., 2011). Despite this, many people are taking drugs that will not benefit them. In a recent survey of the top ten highest selling drugs in the USA, it was reported that for each person benefitting from any of these drugs, between 4 and 25 people are not being helped (Schork, 2015). The healthcare sector is therefore in need of transformation, both to reduce costs and to ensure better treatment of patients. This requires that physicians consider the large variation between individuals to reach this objective. Most currently used pharmaceuticals have been developed based on clinical trials involving large cohorts and are given to patients on the assumption that everyone will respond similarly. This is neglecting the fact that there are large genetic and environmental differences between individuals, and recently it has also been found that the gut microbiome has an influence on drug response, adding further complexity. In order to take these variations into consideration, there is increasing interest in the concept of personalized medicine, which is based on stratification of patients into different molecularly defined groups and then using different treatments and/or interventions for each group (Figure 1A). With accumulating knowledge on the molecular mechanisms for many diseases and the development of more efficient diagnosis, there is increasing interest in moving from disease treatment, the current practice, to disease prevention, as this will significantly reduce costs in the healthcare sector. This, however, requires that identified biomarkers have truly predictive strength, something that can only be obtained through dedicated clinical studies, preferentially longitudinal over long time. Studies that focus on individuals, known as *N*-of-1 trials (Figure 1B), are important for this, as these will provide data on variations within and between individuals and hereby will enable the identification of which biomarkers that can be used for solid stratification and for detection of disease onset. Often *N*-of-1 trials engage the patients actively, i.e., they become participatory. The concept of preventive, predictive, personalized, and participatory medicine has been coined P4 medicine (Hood and Friend, 2011), and this may completely transform the healthcare sector in the next 10–20 years.

Through *N*-of-1 trials on a large number of people, it will become possible to develop more detailed data-driven models for how different biomarkers, and possible even with different quantitative levels, are associated with disease development and thereby enable precision medicine, where the treatment is tailored to deal with a specific molecular event underlying the disease. There are already some ongoing *N*-of-1 studies where healthy people are being monitored in detail over time, one being the so-called 100K project where the objective is to enroll 100,000 individuals and follow them longitudinally for 20 or more years (Hood and Price, 2014). This project follows a 9-month pilot study called the Hundred Person Wellness Project (HPWP), where 100 individuals were intensively monitored and offered regular feedback and counseling about lifestyle changes, e.g., suggested dietary changes or altered sleeping habits (Gibbs, 2014). In the HPWP, all individuals had their genome sequenced at enrollment. Furthermore, insulin sensitivity, immune cell activity, 100 key proteins, and the gut microbiome composition were monitored at enrollment and every 3 months. Finally, pulse and sleeping patterns were monitored continuously using a wrist sensor. This resulted in the generation of a very large amount of data for each individual, a virtual data cloud consisting of billions of data points, and the challenge ahead is to ensure efficient analysis of these data and extract information that can be used for direct advice on lifestyle and/or treatment strategies (Hood, 2013). Analysis of this kind of big data is challenging, as discussed later, but through the integration of the data with mathematical models or reconstructed biological networks, much new biological information can be derived. The use of computational and mathematical models for studying biological systems is referred to as systems biology, and when applied specifically for studying human diseases as systems medicine (Hood, 2013). Here I will discuss the challenges of systems medicine and illustrate how one type of mathematical model,

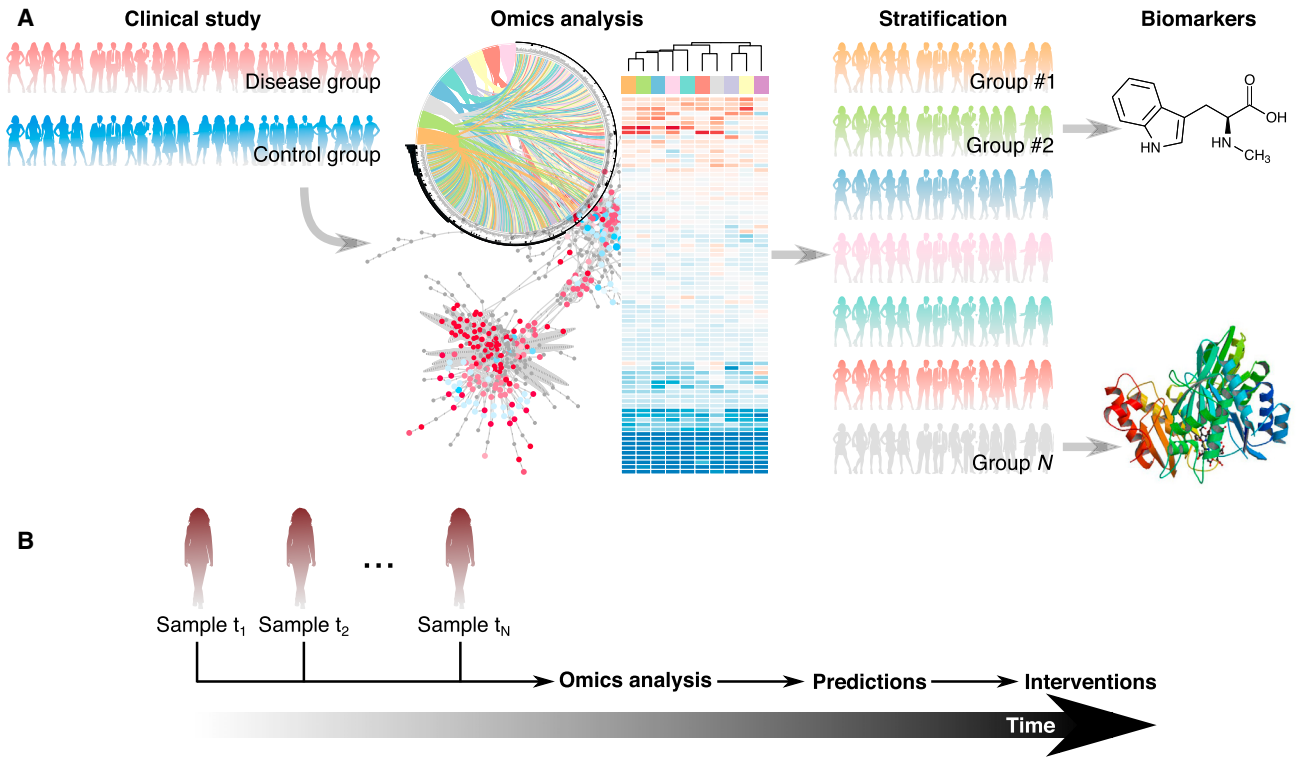

Figure 1. Principles of Personalized Medicine

(A) Illustration of how analysis of big data obtained from detailed omics analysis of patient cohorts can result in detailed phenotyping and thereby lead to stratification of patients into different groups. In connection with this, there can be identified a set of biomarkers that can be used for the stratification in the clinic. These biomarkers are unique molecules (or combination of molecules), e.g., metabolites or proteins, that pass a certain level when specific cellular processes are changed in connection with disease onset or progression.

(B) Illustration of the concept of *N*-of-1 trials. Each individual in the cohort is followed over time, during which samples are taken at different time points. This detailed phenotyping over time enables identification of deviations from normal, which may point to disease development. Furthermore, *N*-of-1 trials will provide information on variations of biomarkers both within and between individuals, and this will be important for identification of biomarkers that are truly predictive and can therefore be used for stratification of cohorts not included in the *N*-of-1 study.

the so-called genome-scale metabolic model (GEM), can be used as a scaffold for integrative analysis with the objective to identify novel prognostic biomarkers that can assist in the advancement toward personalized and precision medicine.

Challenges for Systems Medicine

Advancing systems medicine faces several challenges: (1) the challenge of analyzing large datasets, (2) the difficulties in identifying mechanistic causes for many biomarkers and drug targets, (3) problems with translation from model systems to the clinic, and (4) problems with sample heterogeneity.

The detailed analysis underlying systems medicine results in generation of very large datasets, generally referred to as big data. Even though they are smaller in size than other types of big data generated, e.g., in the financial sector, traffic control, and meteorology, it is challenging to analyze multiple types of omics data as there is a large variation in data structures and formats. Thus, a recent analysis demonstrated that with four different data types, the resources required for data analysis are larger than the resources for data generation for only four datasets, and the resource requirement for data analysis increases rapidly when more datasets are to be analyzed (Palsson and Zengler, 2010). This is because different data types need to be pre-processed separately before they can be used for integrative analysis. Another challenge is that multi-omics data represent varying types of information with very different timescales and different dynamic ranges. Thus, metabolites change with completely different time constants than mRNAs and proteins, and the level of metabolites in a cell is determined not only by the enzyme levels, but also by the kinetics of the individual enzymes; by post-translational modification of enzymes, e.g., protein phosphorylation and acetylation; and by metabolite regulation. Furthermore, metabolites circulating in the blood are determined not only by the metabolic activity of the different tissues, but also by the food intake and by the metabolic activity of the human gut microbiota. Similarly, the plasma proteome is a complex function of the physiological state of the different tissues and cell types in the body. It is therefore challenging to apply plasma metabolomics or proteomics for diagnosis and to integrate these data with other omics data, unless one uses a scaffold that provides a priori information on how the different variables are connected.

Metabolism is closely integrated with practically all cellular processes, and any kind of perturbation in cellular physiology therefore typically results in an altered metabolic footprint, i.e., altered uptake or secretion of metabolites from or to the blood. Plasma metabolomics data therefore have a huge potential for

identification of altered health status. This concept is already widely used for monitoring of triacylglycerides and cholesterols in the blood, but analysis of blood chemistry could probably be used much more widely for diagnosis. However, it is a major challenge to link an altered metabolite profile to onset or progression of a specific disease. To illustrate the complexity, the Human Metabolome Database (HMDB) includes about 42,000 metabolites (Wishart et al., 2013; www.hmdb.ca). A large number of these are food metabolites (about 32,500) and drug metabolites (about 2,500), but still about 4,500 metabolites have been reported in serum (Psychogios et al., 2011; www.hmdb.ca). With so many metabolites present in human serum, and the large sensitivity of the levels of many of these toward lifestyle differences, in particular diet, it is of course challenging to identify biomarkers associated with specific diseases solely from human serum metabolome analysis. Only few biomarkers have therefore been identified from this kind of analysis, exemplified by the identification of elevated levels of branched-chain amino acids as a marker for obesity and diabetes (Newgard et al., 2009; Zhao et al., 2016). Even though transcriptome analysis of abdominal human fat biopsies, enriched in adipocytes, showed that elevated levels of branched-chain amino acids in obese subjects may be caused by reduced respiratory metabolism in this tissue (Mardinoglu et al., 2014a), there is still lacking a mechanistic explanation for why these amino acids are such strong biomarkers for obesity and diabetes. More detailed analysis of metabolic alterations in different tissues is required for obtaining a mechanistic explanation for this finding, and this will require both large datasets, e.g., transcriptome or proteome data, from different tissues in large cohorts, and detailed models that can be used for integrative analysis of such data. The difficulties with identification of prognostic biomarkers solely using plasma metabolomics are well illustrated by identification of sarcosine for prostate cancer progression (Sreekumar et al., 2009). Later studies could not validate these findings (Jentzmik et al., 2011; Ankerst et al., 2015), and like many other biomarkers it has therefore not been translated for clinical use. Through the combination of plasma metabolomics with other omics data, it is possible to get a mechanistic explanation for changes in metabolite levels. This has been demonstrated in several studies where plasma metabolomics was used for genome-wide association studies (GWASs). Using GWASs of more than 200 metabolites in a large cohort of more than 2,000 subjects with a detailed cardiometabolic phenotyping resulted in identification of inborn mutations in *AGXT2*, a transaminase, being associated with altered cholesterol and triacylglycerol levels (Rhee et al., 2013). In a later study on the same cohort, but using more detailed genetic profiling, mutations in several metabolic enzymes were identified to be associated with altered plasma metabolite levels (Rhee et al., 2016). GWAS analysis has also been done with metabolome data from urine samples, and hereby several loci were identified that have also earlier been identified to be important for clinical outcomes, and this led to the identification of several potential metabolite biomarkers that can be measured in urine (Suhre et al., 2011). These are examples of single genetic differences specifically causing an altered enzyme activity, and they are valuable for identification of patients with increased risks for disease development, but GWASs of plasma metabolome data do not allow for gaining insight into metabolic changes associated with disease onset that is not caused by genetic dispositions.

Another challenge for systems medicine is that even though there are many good model systems available for studying different human diseases, the translation to the clinic often fails. This is often ascribed to biological differences between, e.g., mouse and human, but it may equally well be due to impacts of diet and lifestyle, as well as the presence of much larger variations in genetics and gut microbiome composition between human individuals in a clinical trial than in a controlled preclinical study. In the field of cancer, complexity is further added by a large heterogeneity across and within tumors, which even questions the traditional histopathological classification of cancers. *N*-of-1 trials on large cohorts will assist in overcoming some of these challenges as it will allow the identification of the commonalities across a population in connection with disease development, i.e., which are truly conserved biomarkers and associated mechanisms, and which are associated with specific genetic differences and/or lifestyles. Such studies will therefore assist in the identification of prognostic biomarkers that can be used for stratification and for prognosis of disease development.

The Central Role of Metabolism in Cellular Physiology

Metabolism plays a central role in living cells, for it provides the energy and building blocks for cellular growth as well as ensuring protection against external stress factors, e.g., xenobiotics and oxidative stress. Metabolism has evolved to support function of the cell and can roughly be divided into three types: (1) central carbon metabolism, which ensures conversion of carbon and energy sources into free energy, redox power, and precursor metabolites required for biosynthesis; (2) biosynthesis, where precursors are converted into building blocks like amino acids, nucleotides, fatty acids, etc. required for cell growth; and (3) secondary and endogenous metabolism, which is typically highly diverse among cells. Enzymes of the central carbon metabolism are the most catalytically efficient but have evolved to generally be smaller than enzymes of other parts of metabolism (Bar-Even et al., 2011). They are, however, still the most abundant in bacteria, single-cell eukaryal cells like yeast, and human (Liebermeister et al., 2014). In microbes, about half of the proteome is allocated to metabolism, with about 25% being allocated alone to glycolysis, whereas in human this number is lower as a larger fraction of the proteome is allocated to cytoskeleton proteins, chaperones, and the spliceosome (Liebermeister et al., 2014). The high catalytic efficiency, small size, and high abundance of enzymes in the central carbon metabolism are consistent with the central role this part of metabolism is playing in ensuring constant provision of energy, primarily in the form of ATP, in handling electron flows by balancing the co-factors NADH and NADPH, and in providing precursors for cellular growth. Thus, the flux through the central carbon metabolism typically exceeds the flux through other metabolic pathways by a factor 10 or more. With these multiple roles, the central carbon metabolism has to be highly connected with the other parts of metabolism, i.e., in yeast ATP is used in more than 200 out of about 1,500 metabolic reactions, and metabolism therefore forms a highly connected metabolic network (J.N., unpublished data). This means that a perturbation of almost any part of metabolism results in a global response in which a large number of enzymes have to alter their

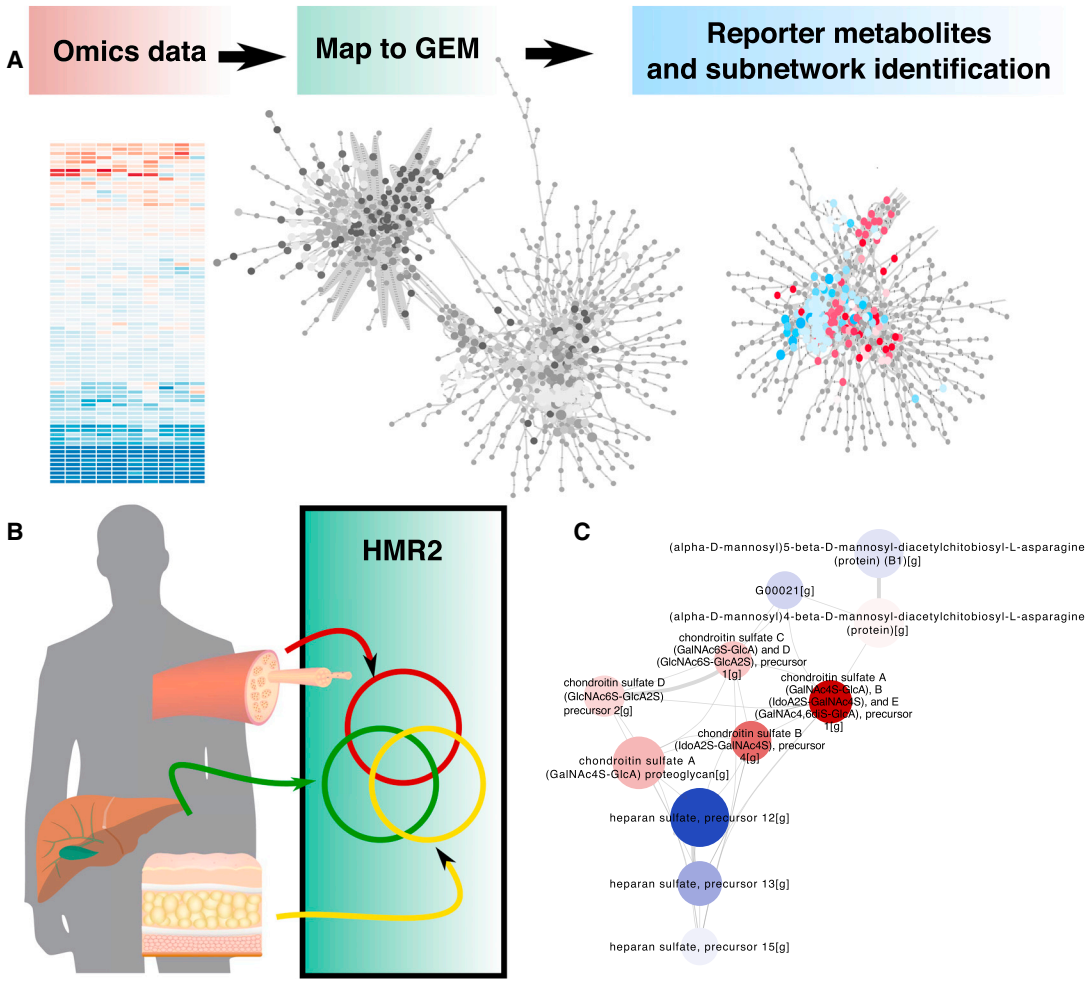

Figure 2. Illustration of the Concept of Integrative Data Analysis Using Metabolic Networks

(A) Illustration of how a metabolic map, represented by a genome-scale metabolic model (GEM), can be used for integrative analysis of omics data, e.g., transcriptome, proteome, or metabolome data. By overlaying these data on the metabolic map, it is possible to identify reporter metabolites and/or sub-networks that represent parts of metabolism that have altered activity in response to change, e.g., disease development. A set of reporter metabolites may be connected in the metabolic network and thereby point to altered activity of non-canonical pathways.

(B) Illustration of how tissue-specific models are a subset of a generic GEM for human metabolism, here illustrated by HMR2.

(C) Example of a reporter sub-network identified in ccRCC using a specific cancer GEM together with transcriptome data from both the cancer tissue and corresponding healthy kidney tissue. The sub-network involves a large number of reactions in heparan and chondroitin sulfate biosynthesis pointing to altered levels of metabolism in plasma and urine.

function in order to maintain homeostasis. This explains why almost any perturbation of cellular physiology will have a metabolic fingerprint, i.e., changes in a certain part of metabolism, and this may be quite specific. It further means that with the high degree of connectivity in metabolism, it is difficult to analyze changes in metabolism without the use of mathematical models. I therefore hypothesize that any disease onset will result in a shift in the metabolic homeostasis in the body, and such shifts can possibly be detected through metabolome analysis of plasma. These changes may be very small, in particular at the early stage of disease onset, and therefore difficult to detect unless a targeted approach is applied. This has to follow a hypothesis generated from analysis of, e.g., transcriptome or proteome data from tissues associated with the disease combined with integrative analysis. As will be discussed below, GEMs represent an excellent scaffold for this kind of analysis.

Genome-scale Metabolic Models

Concept

GEMs are comprehensive compilations of all the metabolic reactions that take place in a particular cell, tissue, organ, or organism (O'Brien et al., 2015). Each reaction is associated with one or more enzymes and encoded by specific genes; thus, a direct gene-protein-reaction connection can be established. This is an important feature of GEMs as it allows for overlaying omics-type data, e.g., transcriptome or proteome data, and thereby identifying co-regulated sub-networks in metabolism (Figure 2A) (Patil and Nielsen, 2005). These co-regulated sub-networks, or reporter metabolites, point to parts of the metabolism that need to have altered expression in order to maintain cellular homeostasis. Often these co-regulated sub-networks are not directly associated with the parts of metabolism that are affected (Patil and Nielsen, 2005). Thus, if cells are exposed to oxidative stress there may be alterations

not only in glutathione metabolism that is directly engaged in coping with the oxidative stress, but also in more distant parts of metabolism, e.g., the pentose phosphate pathway, ensuring regeneration of NADPH used in glutathione metabolism.

Through specification of the stoichiometry of the different reactions in a metabolic network, GEMs can be used for simulation of metabolic functions using the concept of flux balance analysis (O'Brien et al., 2015). This concept assumes that all fluxes into a metabolite pool equal all fluxes out of the pool. Of course, perturbations of metabolism will result in deviations from this steady-state condition, but the flux through most metabolite pools is so high that the pool turnover is on the order of seconds or minutes (depending on the part of metabolism), meaning that a deviation from flux balancing will be resolved in just a few seconds/minutes by the resulting rapid change in metabolite levels. Flux balance analysis imposes a large number of constraints on the fluxes and thereby allows for calculation of fluxes through different parts of the metabolism based on measurements of a few exchange fluxes, e.g., fluxes of nutrient uptake, but as the degrees of freedom in these models is quite large, all fluxes cannot be uniquely determined (Mardinoglu and Nielsen, 2015). Recently it has, however, been shown that by incorporating kinetic information into GEMs, together with a constraint on proteome usage for metabolic enzymes, it is possible to improve the predictive strength of GEMs significantly (Thiele et al., 2012; Nilsson and Nielsen, 2016) and thereby describe overflow metabolism to lactate in cancer cells (Shlomi et al., 2011).

Human GEMs

In 2007, the two first GEMs for human metabolism were reconstructed (Ma et al., 2007; Duarte et al., 2007), and these models formed the basis for Recon2, a much expanded model with broader coverage of metabolism (Thiele et al., 2013). In connection with building tissue-specific GEMs, more details in lipid metabolism had to be incorporated and this resulted in Human Metabolic Reaction (HMR2) (Agren et al., 2014), which is currently the most comprehensive GEM for human cells, covering 3,765 genes, 8,181 reactions, and 6,007 metabolites. HMR2 has been used as a basis for reconstruction of detailed models for different human cell types, which become sub-sets of HMR2 (Figure 2B). Cell-type-specific GEMs have been reconstructed for adipocytes (Mardinoglu et al., 2013), hepatocytes (Mardinoglu et al., 2014b), and myocytes (Väremo et al., 2015). The adipocyte model was used for integrative analysis with the objective of gaining insight into metabolic reprogramming in abdominal fat tissues in response to obesity, and it was found that respiratory metabolism was significantly reduced in obese subjects. At the same time, catabolism of branched-chain amino acids (valine, leucine, and isoleucine) was found to be attenuated (Mardinoglu et al., 2014a), which can explain the elevated levels of these metabolites in plasma (Newgard et al., 2009). The adipocyte model was also used to illustrate that attenuated respiration caused problems with oxidation of accumulated triacylglycerols and therefore resulted in reduced dynamics of lipid bodies in obese subjects (Mardinoglu et al., 2013). The myocyte model was similarly used to identify co-regulated networks in metabolism in response to type 2 diabetes (T2D), and for muscle tissue attenuated catabolism of branched-chain amino acids was identified (Väremo et al., 2015), further pointing to a mechanistic basis for the elevated levels of these metabolites in plasma in obese subjects or those with T2D. Other tissue-specific GEMs have also been reconstructed computationally using data from tissue-specific gene expression values (Shlomi et al., 2008) or from data from the Human Protein Atlas (HPA) (www.proteinatlas.org) (Agren et al., 2012, 2014). HPA data are particularly well suited for the generation of cell-type-specific GEMs, for immunohistochemistry has been used for identifying the presence of proteins in 80 different human cell types, and cell-type-specific models can therefore be generated. These models allow for direct analysis of the metabolism of different cell types present in tissues, and thereby enable better understanding of the mechanisms underlying changes in overall tissue metabolism. RNA sequencing (RNA-seq) has recently been shown to provide much new insight into biological differences between different human tissues, and using this kind of data 32 tissue-specific GEMs were generated (Uhlén et al., 2015). Human GEMs have also been used for the identification of novel drug targets for cancer treatment (Folger et al., 2011), as thoroughly reviewed elsewhere (Yizhak et al., 2015), and recently illustrated for argininosuccinate synthase (ASS1)-deficient tumors (Rabinovich et al., 2015). These tumors have elevated levels of aspartate, which is beneficial for de novo pyrimidine biosynthesis, and it is therefore important to block this part of metabolism in ASS1-deficient tumors. As mentioned above, cancer cells are extremely heterogeneous, and using proteomics data from hepatocellular carcinoma (HCC) tumors, personalized GEMs were generated for six individuals with HCC (Agren et al., 2014). HCC metabolism was indeed found to be quite different in the six individuals, but by using the GEMs it was possible to identify anti-metabolites that block cell growth in all six tumors. One of these targets was the carnitine carrier system, which is responsible for the transport of fatty acids into the mitochondria for β-oxidation and thereby ensures sufficient energy generation for the cancer cells. Using HepG2 cells, a cell line derived from HCC tumors, this target was validated and shown to prevent cell proliferation (Agren et al., 2014). Considering the large heterogeneity in the six tumors, it is, however, very likely that this identified drug target may not constitute an effective treatment across larger cohorts, clearly pointing to the need for a more personalized approach to cancer treatment. GEMs were also used to contextualize gene expression changes independently associated with distinct cancer mutations and revealed a transversal metabolic signature revolving around arachidonic acid and xenobiotic metabolism (Gatto et al., 2016a). This finding may be important as it could lead to the identification of a treatment strategy that can be used for several cancer types.

Identification of Metabolite Biomarkers

GEMs have in several cases demonstrated their power for identification of biomarkers that have subsequently been validated from plasma metabolomics. Using a hepatocyte GEM, it was possible to study metabolic reprogramming in response to development of non-alcoholic fatty liver disease (NAFLD) (Mardinoglu et al., 2014b). From this analysis, it was found that patients developing non-alcoholic steatohepatitis (NASH) had a significant decreased expression of genes encoding for enzymes in serine and glycine biosynthesis, which can explain observation of elevated levels of plasma homocysteine (Gulsen et al., 2005) and decreased levels of phosphatidylserine in the liver of

NASH patients (Gorden et al., 2011). This finding was validated in a follow-up study in which it was shown that NASH patients have reduced levels of serine and glycine in the plasma, pointing to serine deficiency in these patients (Mardinoglu et al., 2016). Moreover, serine supplementation could improve the health status of such patients. This study gives a very strong indication that serine and glycine levels in plasma can be used as a non-invasive biomarker for NASH development in patients with a fatty liver.

HMR2 has also been used to find a very strong prognostic biomarker for clear cell renal cell carcinoma (ccRCC). This was identified from a study that initially evaluated metabolic reprogramming in eight different cancers using RNA-seq data from the Cancer Genome Atlas (TGCA) database (Gatto et al., 2014). From this analysis, ccRCC was found to have a unique metabolic reprograming, distinctive from the other epithelial cancers. This was, in turn, associated with repression of metabolic functions in several different parts of metabolism, e.g., nucleotide metabolism, which makes the tumor more vulnerable against inhibition of specific enzymatic functions according to experimentally validated GEM-based simulations (Gatto et al., 2015). More importantly, the integrative data analysis also identified a strong de-regulation of heparan and chondroitin sulfate biosynthesis, and subsequent quantification of these metabolites in the plasma and urine of patients with metastatic ccRCC resulted in identification of a systems biomarker that is determined by altered levels of several of these metabolites (Gatto et al., 2016b). This systems biomarker was further found to have prognostic value; it can predict the aggressiveness of the tumor and thereby survival rate of ccRCC patients (Gatto et al., 2016c), and it is now being brought to the clinic for evaluation of its diagnostic and predictive capabilities for the treatment of ccRCC.

Finally, a recent study used HMR2 in combination with a biological network derived from protein-protein interactions for analysis of transcriptome and proteome data for insulin-resistant patients and matched controls (Lee et al., 2016). This resulted in the identification of mannose metabolism to be significantly altered in insulin-resistant patients, and subsequent analysis of metabolomics from more than 1,000 subjects could validate mannose as a novel biomarker for insulin resistance (Lee et al., 2016).

The above-mentioned studies are all examples of how systems biology analysis of specific human tissues resulted in the identification of changes in specific parts of the metabolic network, and these changes resulted in altered plasma metabolite levels. It would have been difficult to identify these biomarkers without a directed search, but based on identified and statistically significant alterations in the metabolic networks, a hypothesis could be generated about certain metabolites being likely biomarkers, and from targeted metabolomics these could thereafter be validated. The strength of this approach is that not only are novel biomarkers identified, but a mechanistic explanation for their function is directly provided.

Perspectives

There are some challenges for advancing systems medicine, but these basically condense into developing better methods for integrative analysis of data and the establishment of *N*-of-1 clinical trials with large cohorts. Even though there are several ongoing and planned *N*-of-1 clinical trials, it is important to further expand and include more subjects and also expand the scope of some of these studies to ensure that very detailed phenotypic characterization of the individuals is performed. As discussed, GEMs offer much in terms of integrative analysis, and through further expansion of the models with description of protein synthesis and other cellular processes, the scope of these models will expand and allow for simulating the impact of many key cellular processes underlying human diseases, e.g., oxidative stress and protein mis-folding stress. Other computational approaches should, however, also be considered. Recent development in machine learning, with emphasis on deep learning (Angermueller et al., 2016), has shown to be powerful for analyzing large datasets and holds promise to adapt to problems in computational biology that may in the future assist with diagnostics in the clinic. This was excellently illustrated in a large dietary *N*-of-1 clinical study that was carried out with the objective of enabling personalized dietary advice (Zeevi et al., 2015). Using a very large dataset, involving an 800-person cohort with measured responses to more than 45,000 meals, a machine-learning algorithm was generated by integrating blood chemistry, dietary habits, and gut microbiota composition. Using the algorithm, it was possible to successfully predict glycemic responses in a 100-person follow-up cohort, demonstrating that this algorithm can be used for personalized nutritional advice. Even though machine-learning algorithms cannot directly provide mechanistic insight, these algorithms still allow for providing clear connectivity between a very large number of variables, and these can then be used for follow-up studies with the objective of identifying the underlying mechanisms.

The above-mentioned study, like many other *N*-of-1 clinical trials, included analysis of the gut microbiota, as this has been shown to have a large impact on overall human metabolism (Karlsson et al., 2013a; Arora and Bäckhed, 2016; Wu et al., 2015). However, even though clear correlations have been identified between the gut microbiota and many different human diseases, e.g., T2D (Karlsson et al., 2013b), most of these studies are only correlative and no causal effects have been identified. Here mathematical modeling can assist in gaining insight into the interaction between the many different species and their host (Heinken and Thiele, 2015). The gut microbiota represents a very complex ecosystem with a large number of species that express different metabolic phenotypes. GEMs are well suited for modeling of this kind of ecosystem: models for individual species can capture the overall metabolism of each species, and various algorithms can then be used for simulation of their interactions (Shoaie et al., 2013). Hereby it has been demonstrated that it is possible to simulate how the human gut microbiota is impacted by diet and how it impacts plasma chemistry, including the level of many amino acids (Shoaie et al., 2015). Even though this last study only considered the five most dominant species in the gut microbiota, it clearly demonstrates that it is becoming possible to simulate how this complex ecosystem is impacted by diet and how it interacts with host metabolism. By adding more models, it will become possible to simulate not only the impact of diet on the gut microbiome development but also how the gut microbiome should be modulated, e.g., through addition of new probiotics, in order to attain properties associated with healthy subjects. Here a recent study describing 773

GEMs for gut symbionts provides a valuable resource for expanding our description of the gut microbiota metabolism (Magnúsdóttir et al., 2017). Hereby it may also become possible to use probiotics as combination treatment with drugs that are impacted by the gut microbiota composition, as identified for some anti-cancer drugs (Vétizou et al., 2015; Sivan et al., 2015).

From the above, it is clear that systems biology can lead to identification of novel biomarkers and drug targets, and at the same time provide a mechanistic explanation for why they can be used for diagnosis and in development of effective treatment strategies. However, much more data are needed in order to develop strong biomarkers that are personalized and allow for precise detection of disease onset. GEMs represent an excellent scaffold for analysis of this kind of data, and a particular strength of these models is that they are open ended in the sense that they can be expanded with more biological knowledge and thereby acquire increasing predictive strength. I am therefore confident that together with big data obtained from large *N*-of-1 clinical studies, GEMs will contribute significantly to the advancement of personalized and precision medicine in the next 5–10 years.

ACKNOWLEDGMENTS

I would like to acknowledge valuable discussions with Adil Mardinoglu, Francesco Gatto, and Jon Robinson. I also thank Francesco Gatto with assistance in drafting the figures. I acknowledge funding to my lab from Knut and Alice Wallenberg Foundation, the Novo Nordisk Foundation, Vetenskapsrådet, Bill & Melinda Gates Foundation, FORMAS, and the Swedish Foundation for Strategic Research.

REFERENCES

Agren, R., Bordel, S., Mardinoglu, A., Pornputtapong, N., Nookaew, I., and Nielsen, J. (2012). Reconstruction of genome-scale active metabolic networks for 69 human cell types and 16 cancer types using INIT. PLoS Comput. Biol. *8*, e1002518.

Agren, R., Mardinoglu, A., Asplund, A., Kampf, C., Uhlen, M., and Nielsen, J. (2014). Identification of anticancer drugs for hepatocellular carcinoma through personalized genome-scale metabolic modeling. Mol. Syst. Biol. *10*, 721.

Angermueller, C., Pärnamaa, T., Parts, L., and Stegle, O. (2016). Deep learning for computational biology. Mol. Syst. Biol. *12*, 878.

Ankerst, D.P., Liss, M., Zapata, D., Hoefler, J., Thompson, I.M., and Leach, R.J. (2015). A case control study of sarcosine as an early prostate cancer detection biomarker. BMC Urol. *15*, 99.

Arora, T., and Bäckhed, F. (2016). The gut microbiota and metabolic disease: current understanding and future perspectives. J. Intern. Med. *280*, 339–349.

Bar-Even, A., Noor, E., Savir, Y., Liebermeister, W., Davidi, D., Tawfik, D.S., and Milo, R. (2011). The moderately efficient enzyme: evolutionary and physicochemical trends shaping enzyme parameters. Biochemistry *50*, 4402–4410.

Duarte, N.C., Becker, S.A., Jamshidi, N., Thiele, I., Mo, M.L., Vo, T.D., Srivas, R., and Palsson, B.Ø. (2007). Global reconstruction of the human metabolic network based on genomic and bibliomic data. Proc. Natl. Acad. Sci. USA *104*, 1777–1782.

Folger, O., Jerby, L., Frezza, C., Gottlieb, E., Ruppin, E., and Shlomi, T. (2011). Predicting selective drug targets in cancer through metabolic networks. Mol. Syst. Biol. *7*, 501.

Gatto, F., Nookaew, I., and Nielsen, J. (2014). Chromosome 3p loss of heterozygosity is associated with a unique metabolic network in clear cell renal carcinoma. Proc. Natl. Acad. Sci. USA *111*, E866–E875.

Gatto, F., Miess, H., Schulze, A., and Nielsen, J. (2015). Flux balance analysis predicts essential genes in clear cell renal cell carcinoma metabolism. Sci. Rep. *5*, 10738.

Gatto, F., Schulze, A., and Nielsen, J. (2016a). Systematic analysis reveals that cancer mutations converge on deregulated metabolism of arachidonate and xenobiotics. Cell Rep. *16*, 878–895.

Gatto, F., Volpi, N., Nilsson, H., Nookaew, I., Maruzzo, M., Roma, A., Johansson, M.E., Stierner, U., Lundstam, S., Basso, U., and Nielsen, J. (2016b). Glycosaminoglycan profiling in patients' plasma and urine predicts the occurrence of metastatic clear cell renal cell carcinoma. Cell Rep. *15*, 1822–1836.

Gatto, F., Maruzzo, M., Magro, C., Basso, U., and Nielsen, J. (2016c). Prognostic value of plasma and urine glycosaminoglycan scores in clear cell renal cell carcinoma. Front. Oncol. *6*, 253.

Gibbs, W.W. (2014). Medicine gets up close and personal. Nature *506*, 144–145.

Gorden, D.L., Ivanova, P.T., Myers, D.S., McIntyre, J.O., VanSaun, M.N., Wright, J.K., Matrisian, L.M., and Brown, H.A. (2011). Increased diacylglycerols characterize hepatic lipid changes in progression of human nonalcoholic fatty liver disease; comparison to a murine model. PLoS ONE *6*, e22775.

Gulsen, M., Yesilova, Z., Bagci, S., Uygun, A., Ozcan, A., Ercin, C.N., Erdil, A., Sanisoglu, S.Y., Cakir, E., Ates, Y., et al. (2005). Elevated plasma homocysteine concentrations as a predictor of steatohepatitis in patients with non-alcoholic fatty liver disease. J. Gastroenterol. Hepatol. *20*, 1448–1455.

Heinken, A., and Thiele, I. (2015). Systematic prediction of health-relevant human-microbial co-metabolism through a computational framework. Gut Microbes *6*, 120–130.

Hood, L. (2013). Systems biology and p4 medicine: past, present, and future. Rambam Maimonides Med. J. *4*, e0012.

Hood, L., and Friend, S.H. (2011). Predictive, personalized, preventive, participatory (P4) cancer medicine. Nat. Rev. Clin. Oncol. *8*, 184–187.

Hood, L., and Price, N.D. (2014). Demystifying disease, democratizing health care. Sci. Transl. Med. *6*, 225ed5.

Jentzmik, F., Stephan, C., Lein, M., Miller, K., Kamlage, B., Bethan, B., Kristiansen, G., and Jung, K. (2011). Sarcosine in prostate cancer tissue is not a differential metabolite for prostate cancer aggressiveness and biochemical progression. J. Urol. *185*, 706–711.

Karlsson, F., Tremaroli, V., Nielsen, J., and Bäckhed, F. (2013a). Assessing the human gut microbiota in metabolic diseases. Diabetes *62*, 3341–3349.

Karlsson, F.H., Tremaroli, V., Nookaew, I., Bergström, G., Behre, C.J., Fagerberg, B., Nielsen, J., and Bäckhed, F. (2013b). Gut metagenome in European women with normal, impaired and diabetic glucose control. Nature *498*, 99–103.

Lee, S., Zhang, C., Kilicarslan, M., Piening, B.D., Bjornson, E., Hallström, B.M., Groen, A.K., Ferrannini, E., Laakso, M., Snyder, M., et al. (2016). Integrated network analysis reveals an association between plasma mannose levels and insulin resistance. Cell Metab. *24*, 172–184.

Liebermeister, W., Noor, E., Flamholz, A., Davidi, D., Bernhardt, J., and Milo, R. (2014). Visual account of protein investment in cellular functions. Proc. Natl. Acad. Sci. USA *111*, 8488–8493.

Ma, H., Sorokin, A., Mazein, A., Selkov, A., Selkov, E., Demin, O., and Goryanin, I. (2007). The Edinburgh human metabolic network reconstruction and its functional analysis. Mol. Syst. Biol. *3*, 135.

Magnúsdóttir, S., Heinken, A., Kutt, L., Ravcheev, D.A., Bauer, E., Noronha, A., Greenhalgh, K., Jäger, C., Baginska, J., Wilmes, P., et al. (2017). Generation of genome-scale metabolic reconstructions for 773 members of the human gut microbiota. Nat. Biotechnol. *35*, 81–89.

Mardinoglu, A., and Nielsen, J. (2015). New paradigms for metabolic modeling of human cells. Curr. Opin. Biotechnol. *34*, 91–97.

Mardinoglu, A., Agren, R., Kampf, C., Asplund, A., Nookaew, I., Jacobson, P., Walley, A.J., Froguel, P., Carlsson, L.M., Uhlen, M., and Nielsen, J. (2013). Integration of clinical data with a genome-scale metabolic model of the human adipocyte. Mol. Syst. Biol. *9*, 649.

Mardinoglu, A., Kampf, C., Asplund, A., Fagerberg, L., Hallström, B.M., Edlund, K., Blüher, M., Pontén, F., Uhlen, M., and Nielsen, J. (2014a). Defining the human adipose tissue proteome to reveal metabolic alterations in obesity. J. Proteome Res. *13*, 5106–5119.

Mardinoglu, A., Agren, R., Kampf, C., Asplund, A., Uhlen, M., and Nielsen, J. (2014b). Genome-scale metabolic modelling of hepatocytes reveals serine deficiency in patients with non-alcoholic fatty liver disease. Nat. Commun. *5*, 3083.

Mardinoglu, A., Bjornson, E., Zhang, C., Adiels, M., Barrett, P.H.R., et al. (2016). Personal model-assisted identification of NAD+ and glutathione metabolism as intervention target in NAFLD. Mol. Syst. Biol. Published online March 2, 2017. http://dx.doi.org/10.15252/msb.20167422.

Newgard, C.B., An, J., Bain, J.R., Muehlbauer, M.J., Stevens, R.D., Lien, L.F., Haqq, A.M., Shah, S.H., Arlotto, M., Slentz, C.A., et al. (2009). A branched-chain amino acid-related metabolic signature that differentiates obese and lean humans and contributes to insulin resistance. Cell Metab. *9*, 311–326.

Nilsson, A., and Nielsen, J. (2016). Metabolic trade-offs in yeast are caused by F1F0-ATP synthase. Sci. Rep. *6*, 22264.

O'Brien, E.J., Monk, J.M., and Palsson, B.O. (2015). Using genome-scale models to predict biological capabilities. Cell *161*, 971–987.

Palsson, B., and Zengler, K. (2010). The challenges of integrating multi-omic data sets. Nat. Chem. Biol. *6*, 787–789.

Patil, K.R., and Nielsen, J. (2005). Uncovering transcriptional regulation of metabolism by using metabolic network topology. Proc. Natl. Acad. Sci. USA *102*, 2685–2689.

Pfuntner, A., Wier, L.M., and Elixhauser, A. (2011). Overview of Hospital Stays in the United States. HCUP Statistical Brief #166 (Rockville, USA: Agency for Healthcare Research and Quality).

Psychogios, N., Hau, D.D., Peng, J., Guo, A.C., Mandal, R., Bouatra, S., Sinelnikov, I., Krishnamurthy, R., Eisner, R., Gautam, B., et al. (2011). The human serum metabolome. PLoS ONE *6*, e16957.

Rabinovich, S., Adler, L., Yizhak, K., Sarver, A., Silberman, A., Agron, S., Stettner, N., Sun, Q., Brandis, A., Helbling, D., et al. (2015). Diversion of aspartate in ASS1-deficient tumours fosters de novo pyrimidine synthesis. Nature *527*, 379–383.

Rhee, E.P., Ho, J.E., Chen, M.H., Shen, D., Cheng, S., Larson, M.G., Ghorbani, A., Shi, X., Helenius, I.T., O'Donnell, C.J., et al. (2013). A genome-wide association study of the human metabolome in a community-based cohort. Cell Metab. *18*, 130–143.

Rhee, E.P., Yang, Q., Yu, B., Liu, X., Cheng, S., Deik, A., Pierce, K.A., Bullock, K., Ho, J.E., Levy, D., et al. (2016). An exome array study of the plasma metabolome. Nat. Commun. *7*, 12360.

Schork, N.J. (2015). Personalized medicine: time for one-person trials. Nature *520*, 609–611.

Shlomi, T., Cabili, M.N., Herrgård, M.J., Palsson, B.O., and Ruppin, E. (2008). Network-based prediction of human tissue-specific metabolism. Nat. Biotechnol. *26*, 1003–1010.

Shlomi, T., Benyamini, T., Gottlieb, E., Sharan, R., and Ruppin, E. (2011). Genome-scale metabolic modeling elucidates the role of proliferative adaptation in causing the Warburg effect. PLoS Comput. Biol. *7*, e1002018.

Shoaie, S., Karlsson, F., Mardinoglu, A., Nookaew, I., Bordel, S., and Nielsen, J. (2013). Understanding the interactions between bacteria in the human gut through metabolic modeling. Sci. Rep. *3*, 2532.

Shoaie, S., Ghaffari, P., Kovatcheva-Datchary, P., Mardinoglu, A., Sen, P., Pujos-Guillot, E., de Wouters, T., Juste, C., Rizkalla, S., Chilloux, J., et al.; MICRO-Obes Consortium (2015). Quantifying diet-induced metabolic changes of the human gut microbiome. Cell Metab. *22*, 320–331.

Sivan, A., Corrales, L., Hubert, N., Williams, J.B., Aquino-Michaels, K., Earley, Z.M., Benyamin, F.W., Lei, Y.M., Jabri, B., Alegre, M.L., et al. (2015). Commensal *Bifidobacterium* promotes antitumor immunity and facilitates anti-PD-L1 efficacy. Science *350*, 1084–1089.

Sreekumar, A., Poisson, L.M., Rajendiran, T.M., Khan, A.P., Cao, Q., Yu, J., Laxman, B., Mehra, R., Lonigro, R.J., Li, Y., et al. (2009). Metabolomic profiles delineate potential role for sarcosine in prostate cancer progression. Nature *457*, 910–914.

Suhre, K., Wallaschofski, H., Raffler, J., Friedrich, N., Haring, R., Michael, K., Wasner, C., Krebs, A., Kronenberg, F., Chang, D., et al. (2011). A genome-wide association study of metabolic traits in human urine. Nat. Genet. *43*, 565–569.

Thiele, I., Fleming, R.M., Que, R., Bordbar, A., Diep, D., and Palsson, B.O. (2012). Multiscale modeling of metabolism and macromolecular synthesis in E. coli and its application to the evolution of codon usage. PLoS ONE *7*, e45635.

Thiele, I., Swainston, N., Fleming, R.M.T., Hoppe, A., Sahoo, S., Aurich, M.K., Haraldsdottir, H., Mo, M.L., Rolfsson, O., Stobbe, M.D., et al. (2013). A community-driven global reconstruction of human metabolism. Nat. Biotechnol. *31*, 419–425.

Uhlén, M., Fagerberg, L., Hallström, B.M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, Å., Kampf, C., Sjöstedt, E., Asplund, A., et al. (2015). Proteomics. Tissue-based map of the human proteome. Science *347*, 1260419.

Väremo, L., Scheele, C., Broholm, C., Mardinoglu, A., Kampf, C., Asplund, A., Nookaew, I., Uhlén, M., Pedersen, B.K., and Nielsen, J. (2015). Proteome- and transcriptome-driven reconstruction of the human myocyte metabolic network and its use for identification of markers for diabetes. Cell Rep. *11*, 921–933.

Vétizou, M., Pitt, J.M., Daillère, R., Lepage, P., Waldschmitt, N., Flament, C., Rusakiewicz, S., Routy, B., Roberti, M.P., Duong, C.P., et al. (2015). Anticancer immunotherapy by CTLA-4 blockade relies on the gut microbiota. Science *350*, 1079–1084.

WHO (2011). World Health Statistics 2011 (Geneva, Switzerland: World Health Organization).

Wishart, D.S., Jewison, T., Guo, A.C., Wilson, M., Knox, C., Liu, Y., Djoumbou, Y., Mandal, R., Aziat, F., Dong, E., et al. (2013). HMDB 3.0—The Human Metabolome Database in 2013. Nucleic Acids Res. *41*, D801–D807.

Wu, H., Tremaroli, V., and Bäckhed, F. (2015). Linking microbiota to human diseases: a systems biology perspective. Trends Endocrinol. Metab. *26*, 758–770.

Yizhak, K., Chaneton, B., Gottlieb, E., and Ruppin, E. (2015). Modeling cancer metabolism on a genome scale. Mol. Syst. Biol. *11*, 817.

Zeevi, D., Korem, T., Zmora, N., Israeli, D., Rothschild, D., Weinberger, A., Ben-Yacov, O., Lador, D., Avnit-Sagi, T., Lotan-Pompan, M., et al. (2015). Personalized nutrition by prediction of glycemic responses. Cell *163*, 1079–1094.

Zhao, X., Han, Q., Liu, Y., Sun, C., Gang, X., and Wang, G. (2016). The relationship between branched-chain amino acid related metabolomics signature and insulin resistance: A systematic review. J. Diabetes Res. *2016*, 1–12.